DOCUMENT RESUME

ED 179 763 CE 023 471

AUTHOR Kroll, Arthur M.: Pfister, Linda A.
TITLE Selecting and Using Tests of Career Skills. Information Series No. 177.
INSTITUTION ERIC Clearinghouse on Adult, Career, and Vocational Education, Columbus, Ohio.: Ohio State Univ., Columbus. National Center for Research in Vocational Education.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE 79
CONTRACT 400-76-0122
NOTE 25p.
AVAILABLE FROM National Center Publications, National Center for Research in Vocational Education, The Ohio State University, 1960 Kenny Road, Columbus, Ohio 43210 ($2.20)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Career Education: Evaluation Criteria: *Evaluation Methods: *Job Skills: Measurement: Skills: Test Construction: *Test Interpretation: *Test Reliability: Test Results: Test Reviews: *Tests: Test Selection: Verbal Ability
IDENTIFIERS *Career Skills Measurement: Information Analysis Products

ABSTRACT

The increased attention to measuring career skills has resulted in more instrument development, more testing of students, and more test administrators. There are three key areas of concern. The first area is that of identifying purposes to be served by assessing career skills. Purposes include permitting descriptions of the current status of individuals or groups, and evaluation (effectiveness, utilization in program improvement, allocation of resources, and allocation of money). The second area is that of becoming an informed user of career skills tests--Is the instrument designed for the test audience?: Is the reading level appropriate?: Is the test reliable?: Is it valid?: and Is it free of sexual and cultural stereotyping? The third area concerns limitations in test design and use. There are two principal concerns about the design of paper and pencil career skills assessment materials. They are (1) indirect measurement and the problem of sampling in assessing career skills, and (2) the effects of verbal ability on student performance on measures of career skills. Appropriate caution should be observed in using tests of career skills. (CT)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED179763

Information Series No. 177

SELECTING AND USING TESTS OF CAREER SKILLS

written by

Arthur M. Kroll
Educational Testing Service

and

Linda A. Pfister
College Entrance Examination Board

The ERIC Clearinghouse on Adult, Career, and Vocational Education
The National Center for Research in Vocational Education
The Ohio State University
1960 Kenny Road
Columbus, Ohio 43210

1979

CE 023 471

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE NATIONAL CENTER MISSION STATEMENT

The National Center for Research in Vocational Education's mission is to increase the ability of diverse agencies, institutions, and organizations to solve educational problems relating to individual career planning, preparation, and progression. The National Center fulfills its mission by:

- Generating knowledge through research
- Developing educational programs and products
- Evaluating individual program needs and outcomes
- Installing educational programs and products
- Operating information systems and services
- Conducting leadership development and training programs

FUNDING INFORMATION

Project Title:	ERIC Clearinghouse on Adult, Career, and Vocational Education
Contract Number:	NIE-C-400-76-0122
Educational Act Under Which the Funds were Administered:	Sec. 405 of the General Education Provisions Act, 20 U.S. Code 1221e, (PL 92-318) As amended by PL 93-380 and PL 94-482
Source of Contract:	Department of Health, Education, and Welfare National Institute of Education Washington, D.C.
Contractor:	The National Center for Research in Vocational Education The Ohio State University Columbus, Ohio
Project Director:	Marla Peterson
Disclaimer:	This publication was prepared pursuant to a contract with the National Institute of Education, U.S. Department of Health, Education, and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to freely express their judgment in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official National Institute of Education position or policy.
Discrimination Prohibited:	Title VI of the Civil Rights Act of 1964 states: "No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving federal financial assistance." Title IX of the Education Amendments of 1972 states: "No person in the United States shall, on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any education program or activity receiving federal assistance." The ERIC Clearinghouse Project, like every program or activity receiving financial assistance from the U.S. Department of Health, Education, and Welfare, must comply with these laws.

This publication was prepared with funding from the National Institute of Education, U.S. Department of Health, Education, and Welfare under Contract No. NIE-C-400-76-0122. The opinions expressed in this report do not necessarily reflect the position or policies of NIE or HEW.

FOREWORD

The Educational Resources Information Center on Adult, Career, and Vocational Education (ERIC/CE) is one of sixteen clearinghouses in a nationwide information system that is funded by the National Institute of Education. One of the functions of the Clearinghouse is to interpret the literature that is entered in the ERIC data base. This paper should be of particular interest to practitioners of career education, primarily at the middle school and secondary school levels.

The profession is indebted to Arthur M. Kroll and Linda A. Pfister for their scholarship in the preparation of this paper. Recognition also is due Roger Aubrey, George Peabody College of Vanderbilt University; G. Joely Cowan, American Training and Research Associates; and Donald Fischer, the National Center for Research in Vocational Education, for their critical review of the manuscript prior to its final revision and publication. Robert D. Bhaerman, Assistant Director for Career Education at the ERIC Clearinghouse on Adult, Career, and Vocational Education, coordinated the publication's development. Cathy Thompson assisted in the editing of the manuscript and Bonna Somerlott typed the final draft.

Robert E. Taylor
Executive Director
The National Center for Research
in Vocational Education

ABSTRACT

The increased attention to measuring career skills has resulted in more instrument development, more testing of students, and more test administrators. There are three key areas of concern. The first area is that of identifying purposes to be served by assessing career skills. Purposes include permitting descriptions of the current status of individuals or groups, and evaluation (effectiveness, utilization in program improvement, allocation of resources, and allocation of money). The second area is that of becoming an informed user of career skills tests -- Is the instrument designed for the test audience?; Is the reading level appropriate?; Is the test reliable?; Is it valid?; and Is it free of sexual and cultural stereotyping? The third area concerns limitations in test design and use. There are two principal concerns about the design of paper and pencil career skills assessment materials. They are (1) indirect measurement and the problem of sampling in assessing career skills, and (2) the effects of verbal ability on student performance on measures of career skills. Appropriate caution should be observed in using tests of career skills. (CT)

DESC::*Evaluation Methods; Evaluation Criteria; *Test Interpretation; *Tests; Test Results; Test Reviews; *Job Skills; Test Construction; *Test Reliability; Verbal Ability; Measurement

IDEN::*Career Skills Measurement

CONTENTS

INTRODUCTION	1
PURPOSES SERVED BY ASSESSING CAREER SKILLS	2
BECOMING AN INFORMED CONSUMER OF CAREER SKILLS TESTS	4
LIMITATIONS IN DESIGN AND USE OF TESTS OF CAREER SKILLS	7
INDIRECT MEASUREMENT AND THE PROBLEM OF SAMPLING IN ASSESSING CAREER SKILLS	7
EFFECTS OF VERBAL ABILITY ON PERFORMANCE ON MEASURES OF CAREER SKILLS	12
CAUTIONS AGAINST THE POSSIBLE MISUSE OF TESTS OF CAREER SKILLS	13
SUMMARY	16
REFERENCES	17

INTRODUCTION

The process of assessing career skills has received particular attention in the 1970s with increased interest in career education and concerns about accountability. With the expansion of the importance of career education -- as well as federal and state funds for financial support -- came the requirement for specific ways to evaluate the impact of career education on students. Pressure began to mount for a clearer definition of career education. As Director of USOE's Office of Career Education, Kenneth B. Hoyt provided a definition by advocating a results-oriented approach outlining specific learner outcomes. He states:

> Career education seeks to produce individuals who, when they leave school (at any age or at any level) are
>
> - competent in the basic academic skills required for adaptability in our rapidly changing society;
> - equipped with good work habits;
> - equipped with a personally meaningful set of work values that foster in them a desire to work;
> - equipped with career decision-making skills, and job-getting skills;
> - equipped with a degree of self-understanding and understanding of educational-vocational opportunities sufficient for making sound career decisions;

- aware of means available to them for continuing and recurrent education;

- either placed or actively seeking placement in a paid occupation, in further education, or in a vocation consistent with their current career decisions;

- actively seeking to find meaning and meaningfulness through work in productive use of leisure time;

- aware of means available to themselves for changing career options -- of societal and personal constraints impinging on career alternatives.
(Hoyt, 1977, p. 35)

In essence, the career education movement has progressed to the point that a set of specific skills has been established -- skills that can be taught, learned, and assessed. We refer to these skills as career skills.

The increased attention to measuring career skills has resulted in more instrument development, more testing of students, and, of course, more test administrators. The purpose of this paper is to assist practitioners who select and use career skill tests. Three major topics are covered: identifying purposes to be served by assessing career skills, points to keep in mind if you are to become an informed user of tests of career skills, and limitations in test design with cautions against possible misuse.

PURPOSES SERVED BY ASSESSING CAREER SKILLS

It is good practice to have your purposes clearly in mind when you administer any test of career skills. Such purposes may be as varied as facilitating student development, evaluating program effectiveness, or obtaining base-line data for an experimental research project. If you haven't identified your

anticipated uses of the test results, you should delay administration of the materials until you have planned more carefully. Knowing why you are administering the tests, what you hope to achieve in terms of student development or new information, and how you will follow through in use of the results are important elements in planning.

Why try to assess career skills? What possible purposes can be served by such testing? Four such purposes are among the broad goals of educational assessment in general: first, to find out where individuals stand in their development of career skills and knowledge; second, to identify areas where instruction or guidance may be needed; third, to find out whether such intervention is beneficial; and, fourth, to compare the effectiveness of different forms of instruction.

The first purpose is primarily descriptive. The creation of a yardstick for assessing career skills permits description of the current status of individuals or groups. From such descriptions, the user of the yardstick can evaluate subjectively the quality of different levels of performance.

Once a measure is used and an evaluation is made, a decision can be reached as to whether or not intervention is warranted in the skill development process. This decision, too, is subjective. Although the measure yields data, decisions that involve interpretation of the data rest with the user.

The decision to intervene raises diagnostic questions. What areas of relative strength and weakness have been identified? Which individuals need help and what kind of help? How do the needs relate to possible teaching methods? Can learning experiences be designed to address individual student needs?

The other purpose involves evaluation. How effective are particular instructional strategies? Can assessment of them improve a career education program? Can direction be changed midway to head toward a more promising path? Can the relative effectiveness of different educational programs be determined? Are students in the program developing different or better skills than students who are not? Is the instructional program worth its costs in time and other resources? Should available resources be allocated differently?

Throughout their school careers, students make many decisions, and still other decisions are made by school staff about them.

It often is said that the better the information on which a decision is based, the more satisfactory the decision is likely to be. Tests of career skills cannot make any of these decisions; they can contribute, however, to the information base on which the decisions are made.

BECOMING AN INFORMED CONSUMER OF CAREER SKILLS TESTS

Once you have determined the purposes for which you plan to use career skills tests, you can take a critical look at the assessment tools that are available. Sample sets of instruments are almost always available from commercial publishers. In addition, two publications currently in press which will provide information specifically on career skills tests are: *Career Education Measures: A Compendium of Evaluation Instruments* (McCaslin et al., in press) and *A Sourcebook on Evaluating Career Education* (College Entrance Examination Board, in press).

Get well acquainted with the tests you are considering so that you will understand the purposes for which they were designed. Be sure that the two sets of purposes do not conflict. Whether or not your expertise is in the area of measurement, you can become an informed consumer of career skills testing materials by keeping the following questions in mind as you evaluate each test:

1. *Is the instrument designed for the audience you plan to test.* Different measures are designed for different populations. Check to see what audience(s) the test has been administered to. How does it compare to your own test population. How did the field trial population perform on the test?

You may decide to use a particular test with a population different from that of the test developer. If you are using the test as an instructional or guidance tool, the age or grade level of the measure's focus may not limit your decision. However, if you plan to use the test as an evaluation tool and the test norms or other technical data as benchmarks to compare student performance, then your audience should closely match that used by the test developer.

2. *Does the test measure the content being taught to the proposed test takers?* As career skills tests are often used to evaluate the effectiveness of career education programs, it is essential that you make sure that the objectives of the measure coincide with the goals of your program. Does it measure the content students are being taught? If possible, work through all the questions in a career skills test as you would expect a student to do. Consider each question carefully. Does each question deal with a topic, problem, or activity that you consider useful or important in career development? Can you tell from study of the item what the author was trying to assess through it? Is the item measuring something that is covered in your school's career education or career guidance program?

3. *Is the reading level appropriate for the audience to be tested?* Readability is always a factor to be considered when selecting career skills tests. After all, you want to assess growth in some area of career development, not find out how well someone reads. Some commercial test developers report readability analyses of their material in their technical manuals. For those that do not, you may wish to have the materials examined by a reading specialist in your agency. Remember that the vocabulary of career development -- e.g., apprenticeship, personnel, *Dictionary of Occupational Titles*, *Occupational Outlook Handbook* -- is replete with somewhat difficult words. Yet, for students to be competent in career skills, many people believe that they should have an understanding of these terms. The best test of the effectiveness of an instrument is to administer it to a sample of the students you plan to test. See how they perform and ask for their feedback.

4. *What is the test's reliability?* Data on the test's reliability should be provided in the technical manual produced by the test developer. The reliability of a measure may be defined as the extent to which it yields consistently similar results. Reliability may be expressed as a correlation coefficient, which indicates the extent to which an individual would achieve the same score from one administration of a particular measure to another. If a measure was administered to a group of individuals, the same measure given to the same group several days later, and the two sets of scores correlated, the resulting correlation coefficient would be an index of the measure's reliability. The reliability coefficient is expressed as a numerical value between 0 and 1, although in practice it is usually between .65 and .95 for measures of individual performance. The closer the reliability coefficient is to 1, the more reliable the measure.

Often, in practice reliability coefficients are not derived by a test-retest technique because it is not possible to overcome the effects of practice on the second administration, and for other practical reasons. Instead, the reliability coefficient is usually computed by analyzing statistics for individual questions from a single administration of the instrument; this method yields an index of internal consistency. Such indexes answer the question: To what degree are the questions on the measure interrelated? A high relationship suggests that most of the questions are measuring the same trait or the same combination of traits.

5. *What is the test's validity?* Validity considerations, too, are important in making decisions about whether or not to use a career skills test. A measure may be used in many different ways, and validity as a concept pertains to use of a measure rather than to the measure itself. Some measures may be used for describing how much a student has learned, for identifying students for future participation in a particular skill-building experience, or for deciding which of several instructional approaches designed to develop career skills is most effective. Because each of these applications is based on a different interpretation, the evidence that justifies one application may have little relation to another. Validation thus calls for an integration of many types of evidence.

Three types of validity may be distinguished: content, construct, and criterion-related. Content validity is evaluated by showing how well the content of the measure samples the subject matter or situations about which conclusions are to be drawn. Construct validity is evaluated by determining the degree to which certain explanatory concepts account for performance on the measure. Criterion-related validity is evaluated by comparing scores, or predictions made from them, with another variable considered to provide a direct measure of the characteristic or behavior of interest.

Find out what types of validity studies have been conducted using the test. You may decide to conduct your own content validation to decide the degree to which the test is appropriate for th your particular purposes.

6. *Is the test free of sex and cultural stereotyping?* Goals in career education have stressed the inclusion of all individuals in program efforts. You should examine the materials carefully to determine both obvious and subtle discriminatory statements. Do the items stereotype either sex into any particular roles? Do they represent a multiplicity of cultural backgrounds?

In interpreting test results, be sensitive to the fact that the performance on English-language measures of students whose native language is other than English may be negatively influenced, especially by low reading ability or by lack of familiarity with certain career-related terminology. Similarly, test results are often reported by sex. In general, it makes little sense to compare female and male group performance unless the results are to be used in planning specific skill-building experiences or in conducting of a research study.

LIMITATIONS IN DESIGN AND USE OF TESTS OF CAREER SKILLS

In the history of educational measurement, tests of career skills are relative newcomers. The various commercially available tests vary considerably in their theoretical and conceptual foundations and remain imperfect instruments that must be used with full recognition of their limitations in design and intended use. Two principal concerns about the design of paper-and-pencil career skills assessment materials are: (1) indirect measurement and the problem of sampling in assessing career skills; and (2) the effects of verbal ability on student performance on measures of career skills.

INDIRECT MEASUREMENT AND THE PROBLEM OF SAMPLING IN ASSESSING CAREER SKILLS

When standardized tests are used in an attempt to assess abstract characteristics like career skills, it is important to have a clear appreciation of precisely what is being attempted. Such intangible elements of cognitive functioning cannot be measured directly by "taking a reading" on some standard measuring device and recording the observed measurement. Each test designed to assess career skills is an *indirect* measure that involves a series of readings (one for every item in the test) of various indicators of the skill or knowledge to be measured. The measure is "indirect" in the sense that it assesses a person's responses to a set of structured exercises rather than directly assessing the person's behavior in real-life situations. Something about the skill itself must then be inferred from the information that has been collected from the indicators of that skill.

Many things cannot be measured directly. When we want to know how cold it is outdoors, we do not measure the cold (or heat) itself. We know that heat and cold make things expand and contract. So we use a device like a column of mercury in a thermometer and observe its position and fluctuation to infer how cold it is. Such things as heat, light, electricity, attitudes, emotions, or abilities cannot be measured directly. The presence or absence of such things, or their amount or quality, is determined by inference.

> We measure the effect that these things have on something else. A person's honesty, or aggressiveness, or loneliness could not be measured, but we can infer their presence, or quantity, by observing specific reactions and behaviors which we know are related to honesty, or aggressiveness, or loneliness.
>
> Indirect measurement is a perfectly acceptable procedure and may be seen in fields other than education. For example, the amount of acid in a substance may be measured by the amount of base required to neutralize the acid. Using a similar indirect process, we can measure a pupil's intelligence by observing certain kinds of behavior which we assume are the result of the development of his mental abilities. For example, we might estimate the development of a pupil's memory by observing his ability to memorize. We might estimate a pupil's reasoning ability by observing his performance in solving problems, making generalizations, and discriminations.
> (*Using Tests in the Schools*, 1962, p.5)

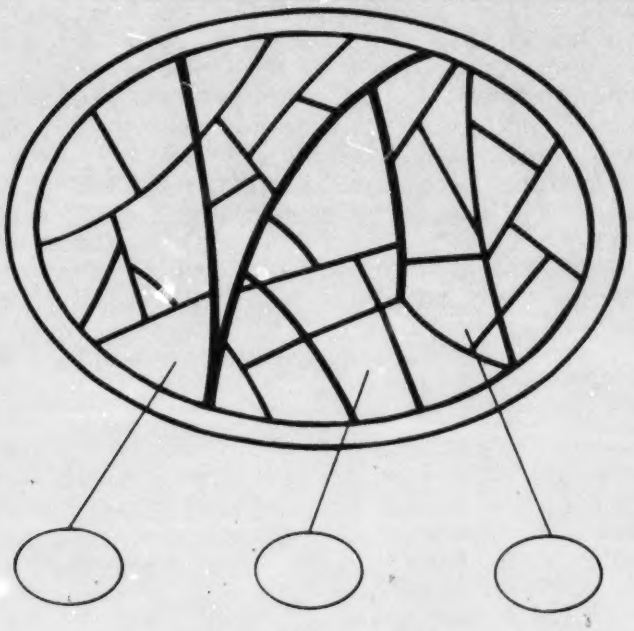
In a way, a career skills measure is a bit like radar, through which observations of a series of "blips" on a screen are used to infer the characteristics of some unseen object. Both are examples of indirect measurement which involves the interpretation of observations. However, the difference between radar and career skills measurement is very great. Radar is applied to physical objects or phenomena where clearly established principles of physical science permit very accurate interpretation on the observations, whereas career skills measurement is applied to cognitive processes where the principles are not yet clearly established,and interpretations are still probabilistic.

Such indirect measurement places a considerable burden on the set of paper-and-pencil test questions from which an individual's career skills and knowledge must be inferred. If the questions or excercises are representative of the universe of career skills and knowledge, then performance on them will yield accurate and useful information. However, to the extent that the exercises are not representative, any inference drawn from the performance data about an individual's career skills may be subject to error. How representative of a universe of career behavior are the items that comprise a given test? Is individual performance on a given sample from the universe of possible items indicative of that individual's potential performance on other item samples drawn from the hypothetical pool of items that comprise the content domain? Can one generalize about an individual's competence on the basis of a given sampling?

Each set of questions that comprises one of the commercially-available measures of career skills assess only a sample of the intricate complex of abilities and functions which are commonly included under the term "career development." In many situations it is both difficult and impractical to assess or evaluate things in their entirety. As a consequence, most evaluations are based on samples. It would be impossible, for example, to test for pollution all of the water in a city's water supply. A few small samples are taken and tested, and the result is used to estimate the quality of all the water.

Many "tests" of this kind have become commonplace. Blood tests are conducted using samples of only a few drops. A winemaster judges the quality of thousands of liters of wine by periodic tasting. A manufacturer of prescription drugs evaluates the content and purity of a product by withdrawing and testing the chemical composition of only a small sample of the continuing output. The butterfat content of a dairy farmer's milk is determined by taking a test tube sample of milk withdrawn from the farmer's refrigerated storage tanks. In each case, the quality of the total amount is estimated on the basis of the sample.

The same process can be used to assess career skills. From the entire complex of skills and knowledge, we evaluate some tiny samples, and from the results of such measurements we attempt to estimate an individual's functional competency in career-related endeavors. The following figure illustrates the process of sampling from the total universe of career skills.



The adaptation of Lewin's (1936) topological representations of psychological structures presented in the illustration is useful in understanding the nature of career skills. The major subdivisions of the career skills universe may be indicated by the thick lines. Some areas are less clearly differentiated, since the skills involved may be shared. Samples may be drawn from a skills subdivision, even though each of these selected skills is so comprehensive that we cannot hope to assess all of it. So, in turn, each subdivision is divided into smaller skill areas, and an attempt is made to sample from each. As a consequence, any career skills test can actually assess only a few samples from the various sectors of a person's total universe of career-related skills.

Sampling problems arise not only from the imperfectly defined and described universe from which they are drawn, but also from inability within a relatively short measure to sample both widely across skill areas and in depth within a skill area. Cronbach and Glaser (1957) discussed this sampling dilemma in the communication engineering terms of bandwidth and fidelity.

> The test designer and the user of tests frequently have to choose between careful estimation of a single variable and more cursory exploration of many separate variables. Tests may be constructed to yield separate scores on a number of diverse, internally homogeneous scales, or to provide a single measure loaded with the general factor underlying items... The dilemma may be described in the language of the communications engineer as a choice between "wideband" and "narrowband" tests. In using a particular channel, such as a telegraph wire, one may either crowd many messages into a period of time, or give a single message slowly and repetitively. The former, more varied message has greater "bandwidth." The wideband signal transmits more information, but the clarity or dependability of the information received is less than for the narrowband signal except under ideal communication conditions. Random errors can seriously confuse the wideband signal; this is spoken of as a lack of fidelity. The tester's situation is analogous. If he concentrated on facts relevant to a single decision, he gets a much more dependable answer than if he spreads his effort. But by concentrating, he leaves all his other questions to be answered on the bases of chance alone.
>
> This suggests that in any decision situation there is some ideal compromise between variety of information (bandwidth) and thoroughness of testing to obtain more certain information (fidelity)... (p. 90)

The sampling problem in assessing career skills, then, is one of finding a number of representative tasks that will require display or use of a reasonable variety of the knowledge and skills in question. The sample must be drawn through some purposive, meaningful procedure, and the sampling process must be described with sufficient thoroughness to permit users of the measures to estimate how adequately performance on the sample tasks reflects performance on all possible career skills tasks.

EFFECTS OF VERBAL ABILITY ON PERFORMANCE ON MEASURES OF CAREER SKILLS

Its degree of validity is usually considered the single most important characteristic of a test. Does the test measure what it purports to measure? Answering such a question for measures of career skills is complicated not only by problems of indirect measurement and sampling, as discussed above, but also by the fact that current skills measures involve reading ability. It is not uncommon for critics of various types of career development measures to comment that they are in fact only measures of reading or verbal ability. However, such a conception assumes that reading is a well-understood process involving certain cognitive activities, the functions of which are separate and unique from other aspects of cognitive activity. This point of view ignores the possibility that the reasoning processes that make reading possible may be the same reasoning processes involved in making career decisions and solving other generally confronted problems of life.

R.L. Thorndike (1973-1974) lent support for consideration of "reading as reasoning" in an article by that title. Thorndike developed evidence that suggested that "performance in reading, at least after the basic decoding skills are mastered, is primarily an indicator of the general level of the individual's thinking and reasoning processes rather than a set of distinct and specialized skills." The three lines of evidence he presented were based on "(1) factorial analyses of specific reading tasks, (2) the correlations between reading tasks and both measures of general intelligence and measuring of later academic progress, and (3) the stability of difficulty in reading test items under translations from one language to another" (p. 133).

He cited an article prepared a half century earlier by E.L. Thorndike (1917) in which the following was asserted:

> Reading is a very elaborate procedure, involving a weighing of each of many elements in a sentence, their organization in the proper relation one to another, the selection of certain of their connotations and the rejection of others, and the cooperation of many forces to determine final responses. In fact... the act of answering simple questions about a simple paragraph...includes all the features characteristic of typical reasoning. (p. 323)

Career skills are no more distinct and separate from general intellectual functioning than in reading.

CAUTIONS AGAINST THE POSSIBLE MISUSE OF TESTS OF CAREER SKILLS

It seems clear that tests of "career skills" are not viewed with sufficient caution. As students wrestle with the eternal question, "What shall I be?", teachers and counselors are forever searching for the magic tool that will help link individuals and opportunities. Yet measurement for purposes of career guidance has made only modest progress over the past 75 years. There continues to be little evidence that differential aptitude tests or interest inventories can provide an adequate basis for helping select among occupational alternatives or for predicting degree of success in an occupation.

Thus, appropriate caution should be observed in using tests of career skills. Users of such materials should avoid the following inappropriate uses of them:

(1) Never use tests of career skills alone as a basis for making predictions about the future career behavior of individuals. Paper-and-pencil tests of career-related knowledge and skills represent at best an imprecise sampling from an uncertain universe of possible knowledge and behavior. Furthermore, such tests should be regarded as status reports that help assess present levels of knowledge and skill. Because knowledge can be acquired and skills developed, what one knows and can do today is of limited use in predicting what one will know or be able to do in the future.

Further complicating use of such tests in predicting future behavior is the uncertain relationship between knowing what to do and doing it. People who behave inappropriately often do so for reasons other than lack of knowledge.

(2) Never assume that a test of career skills is measuring precisely what its title claims to be measuring. A test is considered valid only if it is measuring what it purports to measure. The constructs employed in career development theory are only partially defined, and "career skills" as such probably consist of many such constructs, intertwined in as yet unknown ways. For example, career skills are probably related to general ability. Yet we know little about this relationship.

Because most tests of career skills require reading ability, performance on such measures is likely to reflect verbal ability in large part. In fact, where studies have been done, performance on measures of career skills has correlated at moderate to high levels with performance on verbal ability measures. Correlation studies, however, simply point out degree of association; they should not lead to the conclusion that correlated skills are identical. As one colleague points out, shoe size and foot length are also highly correlated, but shoes are not feet.

Virtually all tests are contaminated by factors other than those they were intended to measure. Paper-and-pencil tests of career skills require the student to read the questions. Poor readers will perform poorly on such tests regardless of their true career skills. The reading ability factor is probably the single most important contaminant of career skills tests.

(3) Be cautious in the use of norm-referenced data in interpreting the results of a career skills test administration. Data about the performance of a group on a career skills test reveal only how the group performed; they tell nothing about how the group *should* perform. Normative data help to describe in relative terms; they do not evaluate.

Brown (1976) noted this common failure to distinguish clearly between measurement and evaluation:

> Measurement answers the question: how much? That is, measurement provides a description of a person's performance; it says nothing about the worth or value of the performance. However, when we interpret a person's performance, we usually place some value or worth on it. At this point we are going beyond description. We are attempting to answer the question: how good? This is evaluation. (p. 12)

Be particularly cautious in interpreting your school's performance on a given test of career skills against any tables of so-called national norms. National norms for a competency area like career skills are difficult to interpret because the schools selected for the norming sample may vary considerably in their emphasis on acquisition of career skills.

(4) Remember that any career skills test provides only a sampling of an individual's skills at a particular point in time. NO test can ever give a completely accurate picture of a person's skills and knowledge. People change. New experiences and learning may result in enhanced skills. The initial sample of questions may have been inappropriate, and the individual might perform quite differently on another sample of questions at a different time.

(5) Be aware of the risk involved in making predictions based on the test scores of an individual. Prediction is usually based on the relationship or degree of association between two sets of data, e.g., scholastic aptitude test scores and freshman grade-point averages. Many studies of this sort report correlation coefficients in the vicinity of .45. According to Franzblau (1958), "...coefficients below .40 do not yield a prediction which is 25 percent better than chance. To yield a prediction which is 25 percent better than a chance or random guess, the correlation must be at least .66" (p. 88). As Barry and Wolf (1962) pointed out, "the inherent errors of the test itself, the inconclusiveness of the relationship between the predictor and what is being predicted make it virtually impossible for the counselor to predict a student's performance with any degree of surety." (pp. 35-36)

(6) Never use the results of performance on career skills measures to evaluate teacher or counselor effectiveness. Performance on educational tests has sometimes been used as a factor in rating the effectiveness of teachers or counselors. We believe that using career skills measures for such a purpose is inappropriate because performance on such measures is influenced by many factors other than staff competency. Students in different classrooms will often have substantially different experiences and characteristics. Parents' educational level, family income, experience in part-time jobs, and the like may influence the attainment of career skills and knowledge. Controlling for such variables in experimental settings is usually quite difficult.

(7) Avoid over-interpreting the results of career skills tests with staff, students, parents, and the community in general. Test scores in general have an aura of exactness about them that is unwarranted. They are simply estimates of what a person can do on a limited sample of exercises or activities at a particular time. Their spurious appearance of being factual needs to be countered by careful explanations of their limitations.

SUMMARY

Changing conceptions of career development have led toward the development of new types of measures for use in career guidance and career education programs. To make effective use of such assessment materials, you should have your purposes clearly in mind and evaluate available materials against those purposes. Finally, remember that available career skills measures are indirect measures that sample imperfectly from the varied universe of knowledge and skills that collectively can be termed "career skills." Because such paper-and-pencil measures require basic reading ability, performance on such measures may be directly influenced by verbal ability. Because such measures possess significant limitations, you should observe appropriate caution in their use and avoid inappropriate interpretation of performance results.

REFERENCES

A Sourcebook on Evaluating Career Education. New York: College Entrance Examination Board, in press.

Barry, R. and Wolf, B. *Epitaph for Vocational Guidance.* New York: Bureau of Publications, Teachers College, Columbia University, 1962.

Brown, F.G. *Principles of Educational and Psychological Testing.* 2d ed. New York: Holt, Rinehart and Winston, 1976.

Committee on Education and Labor. *House Reports: No. 95-150.* Washington, D.C.: House of Representatives (March 31, 1977): 4.

Cronbach, L. and Glaser, G. *Psychological Tests and Personnel Decisions.* 2d ed. Urbana: University of Illinois Press, 1965.

Franzblau, A. *A Primer of Statistics for Non-Statisticians.* New York: Harcourt-Brace, 1958.

Herr, E.L. *Decision-Making and Vocational Development.* Boston: Houghton Mifflin, 1970.

Hoyt, K.B. *A Primer for Career Education.* U.S. Office of Education. Washington, D.C.: U.S. Government Printing Office, 1977. (ED 145 252)

Hoyt, K.B.; Evans, R.N.; Mackin, E.F.; and Mangum, G.L. *Career Education: What It Is and How to Do It.* Salt Lake City, Utah: Olympus Publishing Company, 1972. (ED 062 573)

Lewin, K. *Principles of Topological Psychology.* New York: McGraw-Hill, 1936.

Marland, S.P. *Career Education: A Proposal for Reform.* New York: McGraw-Hill, 1974.

McCaslin, N.L., Gross, C.G., and Walker, J.P. *Career Education Measures: A Compendium of Evaluation Instruments.* Columbus, Ohio: The National Center for Research in Vocational Education, in press. The Ohio State University.

Miller, C.H. "Career Development Theory in Perspective." *Vocational Guidance and Human Development.* Edited by E. Herr. Boston: Houghton Mifflin, 1974.

Parsons, F. *Choosing a Vocation.* Boston: Houghton Mifflin, 1909.

Thorndike, R.L. "Reading as Reasoning." *Reading Research Quarterly.* 9 (1973-1974): 135-145. (EJ 092 582)

Using Tests in the Schools. Los Angeles: Los Angeles County Superintendent of Schools Office, 1962.